

CRITERIA PAPER: ISSUES IN VISIT-BASED BUNDLING

Submitted by:

A. James Lee, Ph.D.
Robert C. Boutwell, M.D., Ph.D.
Janet B. Mitchell, Ph.D.

March 1993

This report was submitted under the Health Care Financing Administration Physician Studies Indefinite Quantity Contract No. 500-89-0050, Delivery Order #4. The views and opinions expressed in this report are the authors' and no endorsement by HCFA or DHHS is intended or should be inferred.

CRITERIA PAPER: ISSUES IN VISIT-BASED BUNDLING

Beginning with the 1984-1986 physician fee freeze and continuing with modest MEI updates, or even fee reductions in the case of overpriced procedures, Medicare has exerted strict control over the rates of increase in Part B payments on a per service basis. Nevertheless, Part B expenditures have continued to increase at double-digit rates (Helbing et al., 1991). The major factor behind this expenditure growth has been increased volume or service intensity. Patients are receiving more services and more complex services than ever before.

In order to effectively control costs in the future, policy makers must control both prices and volumes. The current approach is to employ different tools for the two purposes. The Medicare Fee Schedule is intended to limit prices while the Medicare Volume Performance Standards are intended to control volume. If feasible, a preferred alternative would be to develop payment approaches that simultaneously control both prices and volumes. How might this be done? One promising approach, the one which we explore, is to package or bundle ancillary services with physician visits. Rather than paying for narrowly-defined procedures, payment would be made for a more comprehensive bundle of services related to a specific physician encounter.

The major advantage of such packaging is that it motivates physicians to take a more cost-conscious approach to patient care, giving significant incentive both to reduce or eliminate unnecessary services and also to provide necessary services at lower cost. Under the current system, the physician benefits financially in ordering diagnostic tests whenever he or she also provides those tests, or has an equity position in whoever does. Even when ordering tests from other suppliers with whom he or she has no financial relationship, the physician benefits from the added clinical information without taking any account of the costs.

The current HCPCS procedure coding system includes over 8,000 codes for physician services. The challenge is deciding how to reasonably package or combine those services for payment purposes. A decade ago, researchers at the Center for Health Economics Research (CHER) developed a typology of packaging methods and evaluated a number of specific alternatives using claims data (Mitchell et al., 1983; Mitchell et al., 1987). These alternatives included office visit and surgical procedure packages as well as more comprehensive condition and beneficiary-based packages. At that time, the lack of adequate ambulatory

case-mix measures proved to be a major obstacle in packaging payments for office visits and related ancillaries.

In this study, we give renewed attention to developing and evaluating innovative alternatives for packaging ancillary services with physician visits. Our purpose is not so much to develop a comprehensive framework for visit-based packaging as it is to demonstrate the merit of pursuing additional developmental work in this area.

This paper puts forward a general framework for defining, understanding and evaluating alternative approaches to bundling ancillaries with physician visits. Section 1.0 presents a simplified conceptual representation of the equity-efficiency tradeoff in bundling services together for payment purposes; section 2.0 considers the technical and practical issues involved in visit-based bundling; and section 3.0 recommends criteria or considerations to be investigated in evaluating alternative approaches to visit-based bundling.

1.0 THE CONCEPTUAL CONTEXT

The "ultimate" product being produced by health care providers is the health of the patient. An ideal payment system might pay a fixed amount for a given health outcome. The health care production process, however, is not so well understood that we know how to measure health outcomes or how much to pay for a given improvement in health status. As an alternative to paying for health outcomes, most health care providers are currently paid for the service input they provide to the production process. However, there is no assurance that the inputs provided actually improve health or that they are combined efficiently (i.e., at lowest cost) in producing health improvement. The current reimbursement system thus encourages over-utilization of inputs and does not in any way motivate the selection of lower cost but equally effective treatment alternatives.

The packaging or bundling of services basically involves aggregating inputs to higher levels of intermediate output. The proposed APG classification system, for example, aggregates certain inputs (e.g., ancillaries provided as part of a given visit) into a higher-level input. It is comparable to making the transition from paying for the iron ore to paying for steel when the desired output is a car. An APG-based payment system could encourage more efficient use of inputs in the production of that intermediate output, but it still would not link intermediate outputs directly to health outcomes.

A payment system that provided a fixed dollar amount for the entire episode of illness would more closely link payment to the desired health outcome. With episode-based

payment, the provider would have incentive to provide only the minimum level of services necessary to produce a given health outcome, whether that outcome is to repair damage, prevent further damage, eliminate an infection or alleviate symptoms. The current DRG-based system of hospital reimbursement can be considered a partial episode-based reimbursement system. However, it is not a full episode-based reimbursement system since it does not cover physician services, or pre-admission and post-acute care related to the hospitalization. An episode-based reimbursement system is theoretically appealing but not always practical or feasible.

An additional alternative which links payment to health outcomes is capitation-based payment. A full capitation system pays providers a fixed amount for all health care services provided to an individual or family in a given period of time. A partial capitation system reimburses providers a fixed amount for a category of services provided to an individual or family – e.g., all dental or all mental health services. In theory, capitation reimburses a provider for maintaining a given level of health. The provider has clear financial incentives to do this by using the least-cost combination of inputs.

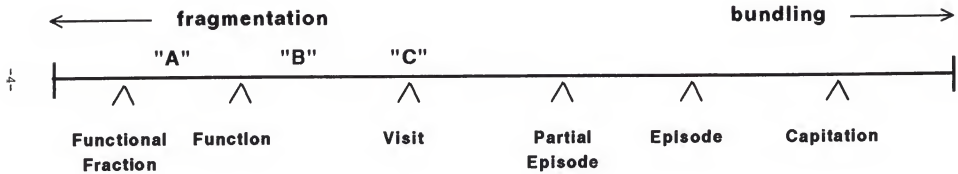
There are, however, two problems with capitation, episode-based payment, or any other bundled payment system. The first is that individuals come to a provider with different levels of health status and thus require varying levels of services to produce a given health outcome. For example, if one of the desired health outcomes is to prevent heart attack or stroke, then a person with high blood pressure requires more services and costs more than a person without high blood pressure. A second problem is that although providers are presumptively being paid for achieving specific health outcomes, there is nevertheless clear financial incentive to provide fewer services and potentially produce substandard outcomes.

1.1 Equity-Efficiency Tradeoffs

Visit-based packaging involves aggregating or combining specific outpatient services currently being paid on an individual, or *a la carte*, basis into bundles of care which would be paid on a single-price, or *prix fixe*, basis. In this sense, it all involves movement to the right along the **packaging continuum** in Exhibit 1. This figure portrays the packaging dimension of health care payment alternatives. As the legend indicates, movement to the left along the packaging continuum implies greater fragmentation (i.e., finer subdivision) of services for payment purposes, and movement to the right implies greater bundling (i.e., grouping together) of services. The relative positions of six basic types of health care

Exhibit 1

THE PACKAGING CONTINUUM



packaging are indicated below the line in Exhibit 1. Beginning on the right-hand side of this figure (i.e., at the highest level of aggregation), we briefly consider what is meant by each of these general types:

Full Capitation. Large medical groups frequently contract with HMOs on a full capitation basis, wherein they are paid a fixed dollar amount per month (or per year) for all professional and institutional services. Less commonly, medical groups contract on a fully-capitated basis for **professional services only**. Inasmuch as we are focusing on professional services, the more restricted definition is appropriate here.

Partial Capitation. IPA-type HMOs routinely contract with primary care physicians on a partial capitation basis, wherein they are paid a fixed dollar amount per month (or per year) for certain, pre-specified professional services. On occasion, other specialties (e.g., mental health) are contracted and paid in this way (Hoy et al., 1991).

Episode. Obstetric services are customarily paid on an episode-of-care basis, that is, including all professional services related to treatment of a specific condition (i.e., pregnancy). Global surgical fees also fall into this category.

Visit. In actual practice, few providers are paid on a pure visit basis, including all services provided as part of or in association with each visit. On an experimental basis, New York State is reimbursing selected outpatient departments and freestanding clinics on a per-visit basis, using its Products of Ambulatory Care (PAC) classification (Moon et al., 1991). A single, fixed payment is made for the professional, facility and ancillary services associated with each clinic visit.¹

Function. Some health care services are billed on a total function basis, namely, including all professional services and supplies related to performance of a discrete treatment or diagnostic test provided to the patient at a single point in time. For example, EKGs were formerly billed on a full function basis (i.e., with a single bill for both the professional and technical components).

Functional Fraction. Much health care is actually paid on a "functional fraction" basis, meaning that patient care functions or services are broken down and billed on the basis of their multiple components (i. e., billing separately for the various inputs involved in providing a treatment or test). Consider two examples. The function or service provided to a patient is a single x-ray, but the professional and technical components can be billed separately. Likewise, the patient receives a single surgery, but the surgeon, assistant surgeon and anesthesiologist are billed separately. These are functional fractions.

Our current CPT-based "fee-for-service" reimbursement system does not fall neatly into any of the above packaging categories. Indeed, as the above discussion suggests, it combines elements of the last four types—namely, (1) episode, (2) visit, (3) function and (4)

¹Ancillary services provided on a later day are also included.

functional fraction-based payment methodologies. We nevertheless suggest that the vast majority of Part B Medicare expense is billed on a functional fraction basis.

For heuristic purposes, our present CPT payment system can be located at point "A" in Exhibit 1, falling midway between the "functional fraction" and "function" categories. By comparison, a payment system based on the newly-proposed Ambulatory Patient Group (APG) classification scheme would be located at point "B," midway between the "function" and "visit" categories. The APG classification does not appear to involve any functional fractions (see Averill *et al.*, 1991). Presumably, any visit-based payment alternatives would fall somewhat nearer to point "C." However, visit-based payment would probably not cover all ancillary services provided in association with every visit. Thus, you would not expect to find any practical payment alternatives actually located at point "C."

As portrayed in Exhibit 2, movement along the packaging continuum involves a fundamental tradeoff between provider equity and social efficiency. A high level of provider equity is assured as long as provider compensation varies directly with the level of work performed; and the more finely the work differences are measured, the greater is the presumptive level of equity. On the other hand, paying providers on the basis of work actually performed gives them incentive to perform more work and provide more services than may be necessary.

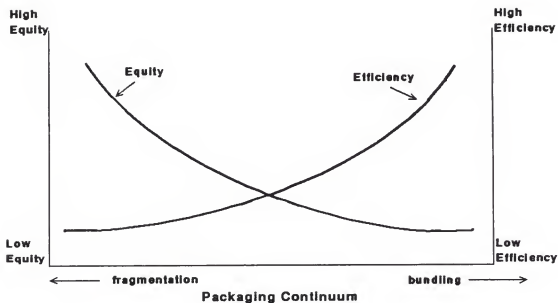
As we move towards the capitated, more highly bundled end of the continuum, the situation reverses. Under capitation, provider compensation varies only with patient volume. That is, the level of compensation for a given patient is altogether independent of the quantity of care actually provided, and significant inequity is possible. A capitated payment approach, however, also gives providers dramatic incentive to avoid unnecessary services and to provide needed services as inexpensively as possible. Thus, movement to the right along the packaging continuum involves a presumptive increase in the overall efficiency of health care delivery.

Our discussion of the equity-efficiency tradeoff is necessarily a simplification of reality. From a policy perspective, determining the "best place" on the packaging continuum requires considering issues that concern access to care and appropriateness of care. For example, provider equity is believed to be correlated with access to care. If providers are not being compensated fairly for the work performed, they have incentive to discriminate and provide care to only the most profitable patients; and the less profitable patients may have difficulty obtaining health care. Like issues of access, there are unanswered questions about how or whether the quality of care varies as one moves along the packaging continuum.

Exhibit 2

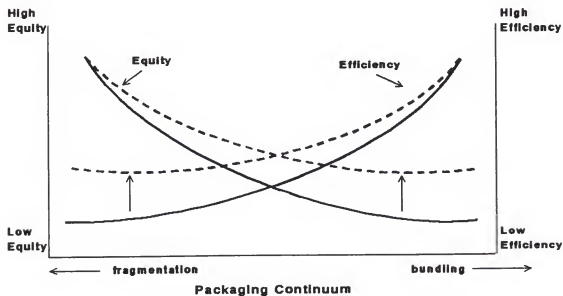
(a)

THE "EQUITY-EFFICIENCY" TRADE OFF



(b)

SHIFTING OF THE "EQUITY-EFFICIENCY" TRADE OFF



Packaging services and capitation-type incentives may cause providers to skimp on necessary care. If so, the cost of health care may be reduced, but the overall efficiency (or appropriateness) of care could also be diminished.

The equity-efficiency tradeoff is not necessarily an immutable relationship, in the sense that the "equity" and "efficiency" levels currently achievable at each point along the packaging continuum can not eventually be improved. As depicted in Exhibit 2, the level of efficiency on the left-hand side of the continuum may perhaps be shifted up (e.g., through the development and application of practice guidelines and various utilization management programs to reduce unnecessary or inappropriate care). Analogously, the level of equity on the right-hand side of the continuum may be shifted up (e.g., through case-mix adjustment, risk pooling, outlier payments and other alternatives for risk-adjusting provider payments to assure that relevant patient acuity differences are taken into account). Health services research and health system reform initiatives are now working to accomplish both objectives, namely, (1) shifting up the equity curve and (2) shifting up the efficiency curve.

Specific strategies have been incorporated into existing payment systems to affect the equity/efficiency tradeoff, as well as movement along the packaging continuum. Examples of such strategies include:

- retroactive payment adjustment (e.g., outlier policies applied to DRG-based hospital payment);
- specification of the bundle definition (e.g., deciding what's included in global reimbursement for surgical services);
- setting payment levels (e.g., determining the conversion factor for the Medicare Fee Schedule); and
- selective exclusions (e.g., not including the "starred" surgical procedures in the global surgery policy).

These strategies reflect important considerations for future reform initiatives.

In this study, we will be primarily concerned with the objective of shifting up the efficiency curve. In an attempt to simultaneously control price and volume, we investigate various alternatives for bundling visits and ancillaries together and thereby moving to the right of the present fee-for-service, *status quo* position on the packaging continuum. In so doing, we shall also develop methodologies for risk-adjusting the packaged payments, and thereby work to minimize the potential adverse impacts on provider equity.

2.0 TECHNICAL AND PRACTICAL ISSUES IN VISIT-BASED BUNDLING

A conceptual understanding, unfortunately, doesn't take one very far in the direction of developing a practical payment system. Operationalizing visit-based payment alternatives requires focusing attention on the technical and other system design issues.

Unit of Payment For this study, the unit of payment is clearly the "visit," and ancillary services are bundled together with the visit for payment purposes. As the APG payment system acknowledges, however, the visit itself is not always the most appropriate unit of payment for ambulatory services. In particular, for some visits, a scheduled "significant procedure" (e.g., upper GI endoscopy) often constitutes the entire reason for the visit and the procedure itself dominates the time and resources expended during the visit. In such instances, it perhaps makes more sense to bundle the visit and related ancillary services together with the procedure (i.e., making the procedure itself the unit of payment).

In order to maintain the focus on visit-based bundling, we may wish to limit this investigation to physician visits that do not involve certain significant procedures. We anticipate that such restriction would be considerably less important for office visits than it might be for OPD visits, inasmuch as major procedures are less commonly performed in office settings. However, the 3M Health Information Systems definition of "significant procedures" (including, as examples, chemotherapy, echocardiography and urinalysis) is much too broad for our purposes. We will formulate our own definition, based in part on our findings from descriptive analysis. In so doing, we will also be guided importantly by findings from an earlier CHER study (Bogen, Boutwell, and Mitchell, 1989).

Bogen, et al studied two types of special procedure packages, one packaging office visits with endoscopies and the other packaging visits with the "starred" surgical procedures. They found little evidence that physicians currently submit bills for office visits when performing endoscopies, with two exceptions. Same-day bills for office visits or consultations accompanied 18 percent of all proctosigmoidoscopies and 10 percent of all sigmoidoscopies. By contrast, visits bills were submitted for only two percent or less of the bronchoscopies, upper GI endoscopies, and colonoscopies beyond the splenic flexure.

The small numbers of visit bills accompanying bronchoscopy, upper GI endoscopy, and colonoscopy were consistent with clinical expectations. These three endoscopies are almost always performed as separate physician services, after patient evaluation and preparation. In almost all instances, they are the sole or primary reason for the physician visit.

For the two less complex endoscopies, however, the greater frequency of billed visits probably reflects the fact that physicians are also providing other services at the same time (e.g., assessment of chronic disease). Attempts to package these endoscopies would nevertheless encourage physicians to unbundle by requesting that patients return at a later date for the endoscopy alone.

The starred surgical procedures are a heterogeneous mix of diagnostic and therapeutic procedures performed by physicians of many different specialties. Examples of high-volume starred surgical procedures are:

- dermatologic procedures, like destruction of skin lesions and cryotherapy;
- orthopedic procedures, like arthrocentesis and injection of tendon sheath; and
- urologic procedures, like catheterization of urethra and dilation of urethral stricture.

Bogen, *et al.*, found that bills for office visits or consultations accompanied one-third to two-thirds of the starred surgical procedures, depending on the type of procedure. There were, however, no consistent patterns by physician specialty or by procedure group (e.g., orthopedic procedures).

Like proctosigmoidoscopy and sigmoidoscopy, some of the starred surgical procedures are performed as part of broader set of services. For example, arthrocentesis (withdrawal of fluid from an inflamed or infected joint) is likely to be performed in the context of a broader evaluation of the patient for the presence of problems such as gout or blood-borne infection. The frequently sudden appearance of a painful, swollen joint and the relatively minor nature of the arthrocentesis procedure increase the likelihood that the physician will obtain the history, examine the patient, order laboratory tests, and perform the arthrocentesis, all within a single visit.

A physician, however, who infrequently performs arthrocentesis may prefer to perform the evaluation and refer the patient to a second physician for the procedure itself. In this latter situation, the second physician would bill only for the starred surgical procedures and not for a visit, if no additional patient evaluation were performed.

Because many of the starred surgical procedures require relatively little time to perform, they may also take place in the context of patient visits in which other significant but unrelated physician services are provided. For example, a physician providing primary

care services to a patient may decide to remove a few skin tags during a visit in which the patient's hypertension is re-evaluated.

The inherent ambiguity involved in defining or distinguishing major special procedures, clearly presents us with an important technical dilemma. Although we do not have an answer at this time, it is a question that we shall have to answer.

Scope of Services Which ancillaries should be included in the bundle? An important consideration here is clinical meaningfulness. Another important consideration is the potential for adverse or undesired behavior. To be realistic, even when normative standards exist about what should or should not be provided during a given physician visit, the physician still has substantial autonomy in deciding what to do. Thus, one must be sure that a visit-based payment system doesn't inadvertently give powerful incentives for withholding medically necessary and appropriate services—or that it pays for services that are not provided. As a practical matter, this means that you probably don't want to include expensive ancillary services, even if customarily provided in association with a physician visit.

If, for example, an oncologist is being paid for a bundled visit that normally involves a \$300 chemotherapy treatment, the actuarial cost of that treatment will be reflected in the visit fee. However, the oncologist then stands to earn a substantial profit if the treatment is not actually provided. In such instance, the provider incentive to withhold treatment may simply be too great.

The financial incentives entailed by visit-based payment can not be so large that physicians have the opportunity to earn "windfall" profits. Even IPA-model HMOs don't include expensive and exotic ancillaries in their capitation payments to primary care physicians.

This suggests one option, to package clinically-meaningful but relatively inexpensive ancillaries only into the visit. Two other possibilities exist. One would be to simply "cap" the amount included in the ancillary-visit bundle; for example, we might include only the first \$50 of any chemotherapy treatments in the oncologist's visit fee. Alternatively, a dollar threshold for ancillaries could be set, beyond which ancillaries would be excluded from the bundle. However, this would give physicians significant incentive to provide or order "out of package" services (i.e., to substitute more expensive alternatives not included in the ancillary-visit bundle).

Case Mix Adjustments Should ancillary-visit bundles be adjusted for case mix? If so, how should this be done? APGs offers one alternative for case mix adjustment.

Although the actual development of new case mix measures is beyond the scope of this project, relevant case mix differences can perhaps be proxied from type of visit (e.g., new vs. established patient) and specialty type. Indeed, it would be extremely convenient (though unlikely) if adequate case mix control could be achieved simply by adjusting the payment rates for existing visit codes.²

Case mix adjustment may be necessary for reasons of both efficiency and equity. It would clearly be inefficient to pay physicians more than what's appropriate, given their patient needs and case mix complexity. However, it would also be inefficient to pay them less than what's appropriate, inasmuch as that could lead to access problems and undertreatment. Also, it would be inequitable for physicians to realize gains and losses unrelated to practice efficiency. That is, you wouldn't want efficient physicians to lose money on bundled payment, and you wouldn't want inefficient physicians to make money on it.

Addressing the exigencies of case mix adjustment does not necessarily require that visit-based payments approximate costs for each individual patient. Whereas that might be a worthy objective, it is almost certainly an unattainable one. Realistically, risk pooling is an essential aspect of any "bundled" payment system. The objective is not so much to pay appropriately for individual patients as it is to pay appropriately for the "average" patient. After all, under bundled payment, providers are not taking risk on individual patients; rather, the continued viability and efficiency of their medical practices depends upon their being appropriately compensated for their entire patient panel. This is a concept that IPA-model HMOs have taken to the extreme in capitating physicians for all primary care services.

Outlier Policies As a practical matter, the domain of impact must be "banded" such that no physicians gain or lose too much in introducing a visit-based payment system. Indeed, this will be one of our principal criteria in selecting among the various payment alternatives and identifying one or more approaches which have the greatest potential for implementation.

²Unfortunately, CPT-4 visit codes do not clearly distinguish visits of different types, and there is a concern that they have been used inconsistently and inappropriately in the past. Studies (e.g., see PPRC, 1989) indicate wide variation in coding practice for evaluation and management services. Although the new visit/consult codes are somewhat more explicit (e.g., indicating expected length of time by visit type), initial provider confusion about how to use these unfamiliar codes could mitigate their value in the near term.

Outlier trimming offers a potentially important mechanism for limiting the provider risk associated with bundled payment. If, for example, the distribution of per visit expense were highly skewed toward low cost, the inclusion of high cost patients in determining the bundle price could result in overpayment for the majority of patients while not substantially reducing the risk entailed by truly "high cost" patients. A second issue, payment rules for outlier cases would need to be specified. Under the DRG system, a percent of the cost beyond the band is reimbursed. Such a policy would both reduce physician risk relative to high cost patients and limit their incentives to provide unnecessary services.

Time Period While the bundle could be limited to those ancillaries provided on the same day as the visit, a longer time period would capture those ancillaries ordered during the visit but not provided until some time later, or those ancillaries performed prior to the visit during which the results will be reviewed. Key issues here include the length of the "time window" surrounding the visit and whether the window should be symmetric (i.e., to include the same number of days before and after the visit).

At present, this question is substantially an empirical one. How many ancillaries would be missed by indicating various alternative time parameters? It is only necessary to be reasonably comprehensive to induce greater efficiency.

In the event that physicians respond to visit-based payment by shifting ancillaries outside the defined "time window," it would then become necessary to consider alternative strategies. For example, HCFA could investigate administrative options for uniquely associating ancillaries with individual physician visits.

Who to Pay Visit-based payment would unquestionably impact on physician relationships with heretofore independent ancillary providers. Indeed, in terms of efficiency, that is a desired consequence of moving towards bundled payment.

To the extent that ancillaries are provided by someone other than the referring physician, multiple providers are involved. If the visit-based payment covered those services provided by the physician only, physicians would have compelling incentive to make ancillary referrals. It seems inescapable that many of the services provided by independent ancillary providers must also be included in the visit fees.

Payment for ancillary providers could take one of two forms. The visit physician could assume responsibility for paying the ancillary providers. Alternatively, HCFA could pay ancillary provider bills directly, as long as the ancillary providers bill at the rates negotiated with the visit physician. Administratively, this would require HCFA to subtract such payments from the visit fees payable to the visit physician. For monitoring and

management purposes, the visit physicians would want reports of the services billed and the amounts paid to ancillary providers on behalf of their patients. Legal constraints would have to be addressed in either case. For example, independent laboratories are already required to bill HCFA directly for their services.

How Much to Pay Should payment be based on current utilization? Except in cases in which generally-accepted clinical guidelines indicate otherwise (e.g., that certain ancillary tests are unnecessary), there is no good alternative but to reflect current medical practice in calculating the visit-based payment fees. Furthermore, absent other policy objectives, we assume that the initial payment rates would be set on a "budget neutral" basis (i.e., not changing the total amount paid for visits and ancillaries but only the way in which they are paid). Without a budget neutral implementation, one could not distinguish the consequences of changing the payment system itself from the concomitant effects of changing total expenditures. Also, it would be difficult to get providers to support permanent change or participate in a demonstration test if it were known a priori that their total revenues would be reduced.

Assignment Although the 15 percent limit on balance billing mitigates the issue, the assignment question is an especially "thorny" one. Visit-based payment may not work if assignment remained voluntary. Consider the following.

If voluntary assignment were permitted, the entire concept of risk pooling (see above) would be undermined. Non participating physicians accepting assignment would limit that patient pool to the easy, inexpensive patient visits, and exclude the more difficult and expensive ones. Medicare would find itself paying physicians considerably more than necessary for assigned visits wherein the bundled price is greater than what Medicare otherwise would have paid (i.e., less expensive visits). Despite balance billing limits, beneficiaries would still pay more out-of-pocket for unassigned visits wherein the bundled price is less than what Medicare otherwise would have paid (i.e., more expensive visits). The result would be an income transfer from beneficiaries to physicians, with Medicare program outlays unaffected.

The answers to questions raised above will be sought from the empirical work conducted in this study. In particular, we will investigate the redistributive and risk consequences of several visit-based payment alternatives (including an APG-like approach). We will then consider whether or not the projected distributional impacts are reasonably equitable and have sufficient potential to warrant the continued development, testing or implementation of such bundled payment arrangements.

3.0 A STRUCTURE FOR EVALUATING ALTERNATIVE APPROACHES

A small multitude of evaluation criteria could conceivably be indicated for assessing the advantages and disadvantages of alternative visit-based payment systems. However, in an effort to simplify and focus evaluative efforts, we indicate only four such criteria, ones which we believe can encompass most important differences between alternative payment approaches:

1. Cost Reduction Potential
2. Distributional Consequences
3. Potential for Inappropriate Responses
4. Administrative Complexity

Each of these is discussed below.

3.1 Cost Reduction Potential

The primary goal in developing visit bundling strategies is to reduce the costs (or, equivalently, the rate of growth in costs) of physician-directed services. The most fundamental difference between the current system and package-based reimbursement is that the link between cost and revenue is severed in the latter. Providers receive a fixed payment for a given package and the payment is altogether independent of the cost of providing the package. Thus profit (Π) on a given patient can be written as:

$$\Pi = R_j - \sum P_{ij} \cdot Q_{ij}$$

where

R_j = reimbursement for package j

P_{ij} = the price of providing service i in package j , and

Q_{ij} = the quantity of each service i in package j

Since the marginal revenue associated with providing an additional service is zero, there is strong incentive to increase profits by reducing costs. Cost reductions may involve either price effects (i.e., changes in the P_{ij} 's), quantity effects (i.e., changes in the Q_{ij} 's), or both.

Package providers can reduce their unit costs by doing either of the following: (1) negotiating lower fees for services obtained from other providers or (2) providing the ancillaries themselves at a lower cost.³ Package quantities can be reduced by providing or ordering fewer ancillaries (i.e., eliminating those of marginal clinical value). It may also be possible to reduce total package costs by changing the service mix and substituting lower-cost services for higher-cost ones included in the package (e.g., substituting a less expensive chemotherapy regimen).

3.2 Distributional Consequences

There are no well-defined, objective criteria for gauging whether the distributional impacts of packaged payment are equitable and acceptable—acceptable both to policy-makers and to the provider community. Nevertheless, three distributional objectives need to be met:

1. The potential for "windfall" gains or losses to particular types of providers or provider subgroups should be minimized.
2. The potential for "windfall" gains or losses on particular types of patients or patient subgroups should be minimized.
3. Compared to the current payment baseline, the dispersion of financial winners and losers should be relatively narrow.

To the extent that these objectives are not met, one confronts three problems or concerns. One, there may not be sufficient homogeneity, in terms of the service frequency or intensity of services included in the visit package. In this case, the *prima facie* validity or integrity of packaging itself may be suspect. Practice variation may reflect provider inefficiency, or it may reflect underlying differences in patient need. If it is the latter, risk pooling on a package basis becomes less tenable. It is extremely difficult, however, to determine whether practice variation is due to practice differences or to case-mix variation. Absent a strong central tendency, there is no mechanism for determining practice standards from claims data alone.

Two, if certain types of patients (e.g., those aged 85 or more) are substantially more or less profitable, providers will have an incentive to seek out the more profitable patient types

³This option assumes that not only the visit physician but also independent ancillary providers are included in the bundled payment.

and deny services to the less profitable. Such provider behavior would not be a desirable outcome.

Three, even if dramatic changes in practice behavior are desired, a package payment system may need to be implemented over a transitional period to allow gradual adjustment to changed payment incentives. Alternatively, for a period of time, one may choose to blend fee-for-service payment with packaged payments.

Providers cannot be expected to adapt or modify their practice patterns all at once. It will take time for physicians to renegotiate or reconfigure their ancillary provider relationships, or to otherwise develop the facility to provide ancillaries more efficiently. Moreover, in the absence of well-established clinical guidelines, physicians will be understandably reluctant to change their practice patterns abruptly. Rather they may implement changes initially on a limited, trial basis and then, as experience proves favorable, gradually extend the changes to their entire practice.

Policy makers may also wish to moderate the bundling incentives at time of initial implementation. They themselves may not have sufficient confidence in bundled payment to demand an abrupt adjustment in practice behavior. A graduated adjustment process would help to reduce the policy risk in this regard.

3.3 Potential for Inappropriate Responses

To the extent that packagable services are currently being provided inefficiently—too high volumes or at prices that are too high—packaging should have the effect of improving practice efficiency. On the other hand, packaging also provides undeniable incentive for providers to compromise the quality of care (e.g., to eliminate necessary and appropriate services, or to substitute an "inferior" mix of services). In implementing any of the package options, one would hope that professional standards and peer oversight will preclude such inimical behavior on a broad scale. Indeed, there is no compelling evidence that the experience of HMOs in capitating primary care physicians and other providers has been otherwise. Nevertheless, one cannot preclude the possibility that some providers will take advantage of the situation and behave irresponsibly. To avoid or minimize such prospects, any demonstration or program implementation must incorporate some mechanism for review

or audit of providers with the lowest package costs (i.e., the low cost outliers).⁴ Access and quality of care for high cost patients (e.g., very old patients with certain comorbidities or chronic illnesses) will also need monitoring.

Perhaps the greatest challenge to implementing visit-based packaging is controlling the **numbers of visits** from Medicare beneficiaries. A provider's package revenues will vary directly with the number of visits, further intensifying existing incentives to see patients more often than necessary. Also, providers would have new incentives to spread out the costs of expensive testing regimens over several visits. In order to avoid abusive expansion in the number of visits, it may be necessary to establish and administer explicit criteria (e.g., practice guidelines) for treatment of specific conditions.

The potential for "out-of-package" billing within visit should be a somewhat less serious concern. A well conceived package should offer limited opportunity for shifting care to other ambulatory care providers who would not be paid within the package. Indeed, in order to avoid such prospects, one must deliberately choose broad package definitions.

Visit-based packages may nevertheless involve significant potential for shifting care into the inpatient hospital setting. Much cardiac testing and cancer treatment, for examples, are done on an inpatient basis, and such inpatient care unquestionably substitutes for ambulatory services. Under a visit-based payment system, the package provider who anticipates an expensive cardiac testing or cancer treatment regimen may be more willing to admit his or her patient to the hospital and thereby possibly shift payment liability back to the Medicare program. Presumably, providers would not take such drastic action for the sake of offloading comparatively small payment liabilities; and an outlier payment methodology could mute the advantage of shifting larger liabilities, by assuring the overall equity of the payment system. In addition, we assume that the PROs (or Provider Review Organizations) will continue their institutional responsibility as hospital gatekeepers, and help to assure that Medicare beneficiaries are not being admitted inappropriately.

Packaging may also create incentives for patient skimming (i.e., providers accepting only less complicated patients). If so, the patients with more complicated problems may have difficulty obtaining access to care. This problem could be avoided if it were feasible to

⁴In order to monitor utilization, it will be necessary to require continued reporting of the services provided, even if those services are being paid on a bundled basis. This is analogous to the situation in many or most HMOs; namely, physicians being paid on a capitation basis still have to submit encounter reports identifying services provided within the capitation.

construct packages in such a way that providers do not know a priori whether or not their patients are good risks. However, to the extent that physicians can screen their patients for risk, the access concern is one which must receive considerable attention in actually designing the administrative procedures for packaged payment.

Like the quality of care concern, the access concern might be met through a plan for targeted review of the lowest cost providers. Also, like HMOs, HCFA might establish a patient grievance process, so that vulnerable, potentially high cost patients who have been denied care could report their experience to someone responsible for provider affairs. Providers who systematically decline to accept more expensive patients might then be penalized or paid on a reduced rate basis.

The extent to which providers respond to packaging incentives in any of the unintended and inappropriate ways strongly affects the desirability of package-based reimbursements.

3.4 Administrative Complexity

The likelihood of actually implementing a visit-based payment system, on either a permanent or trial basis, unquestionably varies inversely with the complexity of the prospective implementation itself. An APG-based payment system, for example, is potentially more difficult and expensive to implement than one that merely adjusts the payment rates for different types of visits and specialties—assuming no outlier policies, windows or bundling across providers.

An APG-like approach may also entail greater policy risk. A payment program failure would be more visible and damaging if providers had been asked to make major changes in their administrative practices (e.g., to buy and implement new software),⁵ and providers subsequently had to reverse or substantially modify those changes. That is not to say that one should not take policy risks of this type, when prudent. In retrospect, it clearly made sense to take a significant policy risk in implementing DRG-based payment for inpatient services.

We can not (and should not) make decisions relative to how much policy risk HCFA should take in implementing or testing alternative visit-based payment systems. Rather, it

⁵Although new software may not be required for billing purposes, hospitals may need it to evaluate and manage their experience with APG payment (e.g., to determine which types of patients are profitable or unprofitable, and why).

will be our role to anticipate and understand, as well as we can, the administrative consequences of alternative approaches. Of course, other things being equal (i.e., ceteris paribus), a less complex implementation is clearly preferred to a more complex one.

The administrative complexity or burden entailed by alternative payment systems should be evaluated along a number of dimensions. One must distinguish the administrative consequences not only for HCFA but also for the carriers and the providers (and possibly even for the patients). To the extent possible, the resource costs entailed by specific administrative requirements should be estimated. In so doing, one would also want to distinguish between the one-time costs of implementation and the continuing costs of operating the payment system. No doubt, the most difficult part of this appraisal will be describing the payment systems with sufficient specificity to adequately anticipate the implementation and operational requirements.

None of these evaluative criteria can be mechanistically applied in profiling the "pros" and "cons" of alternative visit-based payment systems. As we have seen, each involves a variety of behavioral and other considerations; and, for many, their importance can not be reliably ascertained apart from some kind of market or demonstration test. After all, it would have been extremely difficult to foresee the value of capitating primary care physicians (as HMOs do) if no one had ever done it before. We simply don't understand either physician behavior or clinical practice well enough to anticipate the utilization and cost impacts.

REFERENCES

Averill, Richard F., et al Design and Evaluation of a Prospective Payment System for Ambulatory Care, 3M Health Information Systems, Report prepared under HCFA Cooperative Agreement No. 17-C-99369/1-02, December 1990.

Bogen, Jonathan, , Robert C. Boutwell and Mitchell, Janet B. Jonathan Bogen, Packaging Diagnostic Test Interpretation and Surgical Procedures, With Office Visits, HCFA Cooperative Agreement No. 99-C-985261-04, April 1984.

Helbing, Charles, et al "Medicare Expenditures for Physician and Supplier Services, 1970-88," HCFA Review, 12:109-120, Spring, 1991.

Hoy, Elizabeth W., Richard E. Curtis and Thomas Rice, "Change and Growth in Managed Care," Health Affairs, 18-36, Winter 1991.

Mitchell, Janet B., et al Alternative Methods for Describing Physician Services Performed and Billed, Final Report submitted to HCFA, November 1983.

Mitchell, Janet B., Jerry Cromwell, Kathleen A. Calore, and William B. Stason, "Packaging Physician Services: Alternative Approaches to Medicare Part B Reimbursement," Inquiry, Winter 1987.

Moon, Marilyn, Margaret Sulvetta and Mark Miller, "Applying PACs to Medicare: Final Report," Urban Institute Working Paper 3855-02, June 1991.

Physician Payment Review Commission, Annual Report to Congress, 1989, Washington, D.C., 1989.

CMS LIBRARY



3 8095 00012712 2